Stochastic modeling

A few ways to introduce stochasticity:

- Add noise to the duration of events
- Replace reaction rates with reaction probabilities
- Replace deterministic transfer functions with noisy transfer functions
- Add noise to the outcomes

Stochasticity in Boolean models

1. Noise in timescales – implemented using randomness in update orders (seen before)

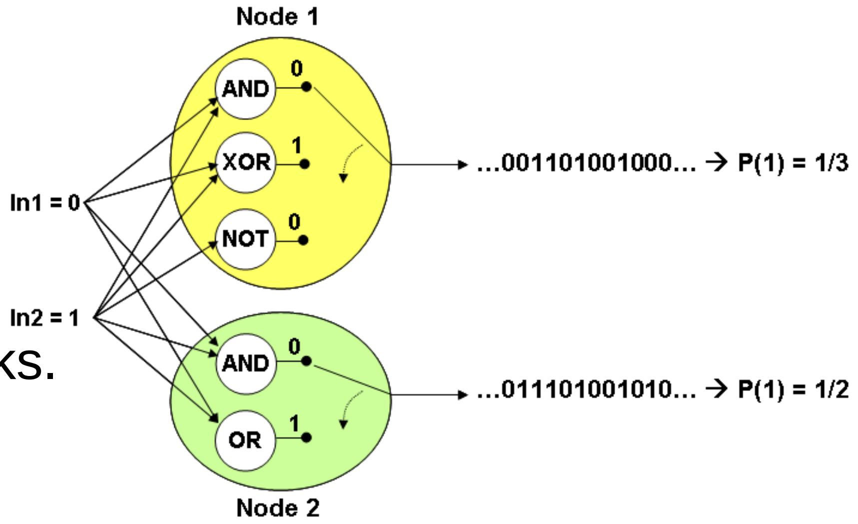
2. Noisy transfer functions

in2	in1	P(out=0)	P(out=1)
0	0	1	0
0	1	p_1	$1 - p_1$
1	0	p_2	$1 - p_2$
1	1	$p_1 \times p_2$	$1 – p_1 \times p_2$

Noisy OR

3. Probabilistic Boolean network: each node makes a probabilistic choice among several transfer functions.

A probabilistic Boolean network represents a stochastic weighted average of multiple Boolean networks.



Bayesian network modeling

- Probabilistic network modeling with relatively few parameters
- Requires large amounts of data, but can use heterogeneous data

- The interactions modeled do not need to be direct
- The basic Bayesian models cannot incorporate feedbacks and dynamics
- Most Bayesian models are discrete, e.g. they assume that the nodes can have one of two states (on, off)

- Main concept: conditional probability, i.e. the probability of an event given that we know some other event has occurred
- Bayesian network models are usually used to find models that fit existing data

An example of a Bayesian network model

$^{\theta}$SI p(signal/stimulant)

Stimulant	High	Medium	Low
Present	0.6	0.3	0.1
Not present	0.1	0.2	0.7

$^{\theta}$IN p(inhibitor/signal)

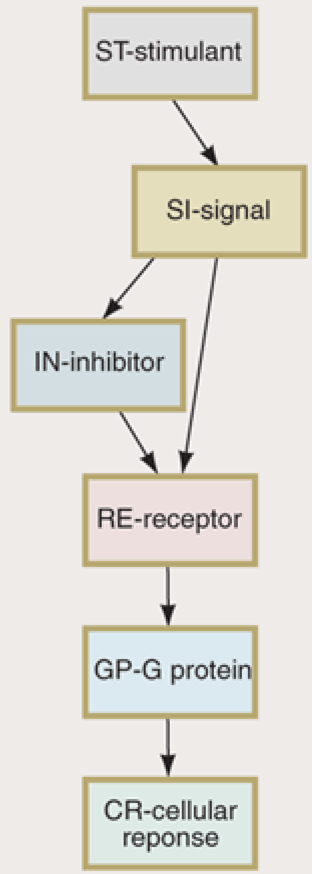
Signal	High	Medium	Low
High	0.6	0.3	0.1
Medium	0.2	0.2	0.6
Low	0.1	0.1	0.8

$^{\theta}$GP p(G protein/receptor)

Receptor binds	Active	Not active
Yes	0.9	0.1
No	0.1	0.9

$^{\theta}$CR p(cellular response/G protein)

G protein	Yes	No
Active	0.8	0.2
Not active	0.1	0.9



$^{\theta}$ST p(stimulant)

	Present	Not present
	0.4	0.6

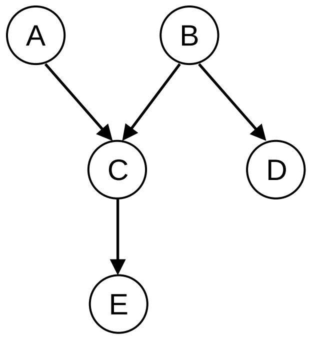
$^{\theta}$RE p(receptor binds/signal, inhibitor)

Signal	Inhibitor	Yes	No
High	High	0.5	0.5
High	Medium	0.8	0.2
High	Low	0.9	0.1
Medium	High	0.3	0.7
Medium	Medium	0.5	0.5
Medium	Low	0.8	0.2
Low	High	0.2	0.8
Low	Medium	0.3	0.7
Low	Low	0.5	0.5

Needham et al Nature Biotechnology 2006

Steps for creating a Bayesian network model

- Summarize/guess the nodes (players). Summarize/guess the interactions/dependency relations, represent them as directed edges. No cycles (feedback loops) are allowed.
- The state of each node is assumed to be determined entirely by its current inputs.
- Assume that the nodes can have one of two states (on, off).
- For each node construct a conditional probability table that gives the probability of the node's states for each combination of states for the regulator nodes.
- Select the state values of source nodes. Use the tables to determine the states of the sink (output) nodes.



A	B	P(C=1)
0	0	0.02
0	1	0.08
1	0	0.06
1	1	0.88

B	P(D=1)
0	0.01
1	0.9

C	P(E=1)
0	0.03
1	0.92

Q: find similarities and differences between Boolean and Bayesian models.

Conditional probability

Joint probability of two variables V, W: P(V=v, W=w). Here v, w are specific values.
Conditional probability: P(V=v| W=w). Means the probability of V=v given that W=w.

P(V=v,W=w) = P(W=w). P(V=v|W=w) = P(V=v). P(W=w|V=v)

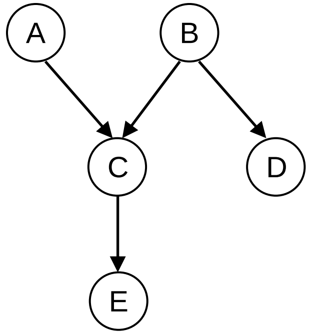
$$P(V=v|W=w)=\frac{P(W=w|V=v)\times P(V=v)}{P(W=w)}$$

Bayes' Rule

To calculate the probability of a node in a particular state we only need to know the states of the nodes in which it is conditioned.

Determining probabilities

P(V=v,W=w) = P(W=w). P(V=v|W=w) = P(V=v). P(W=w|V=v)



The tables indicate conditional probabilities, e.g.

P(C=1|A=0, B=0) = 0.02

P(E=0|C=0) = 0.03

Assume A=B=1.

P(C=1|A=B=1) = 0.88

P(D=1|A=B=1) = P(D=1|B=1) = 0.9

P(E=1|A=B=1) = P(E=1, C=0|A=B=1) + P(E=1, C=1|A=B=1)

=P(C=0|A=B=1)P(E=1|C=0) + P(C=1|A=B=1)P(E=1|C=1)

=0.12*0.03+0.88*0.92 = 0.81

A	B	P(C=1)
0	0	0.02
0	1	0.08
1	0	0.06
1	1	0.88

B	P(D=1)
0	0.01
1	0.9

C	P(E=1)
0	0.03
1	0.92

Q: What is the closest Boolean model? What would it indicate for A=B=1?

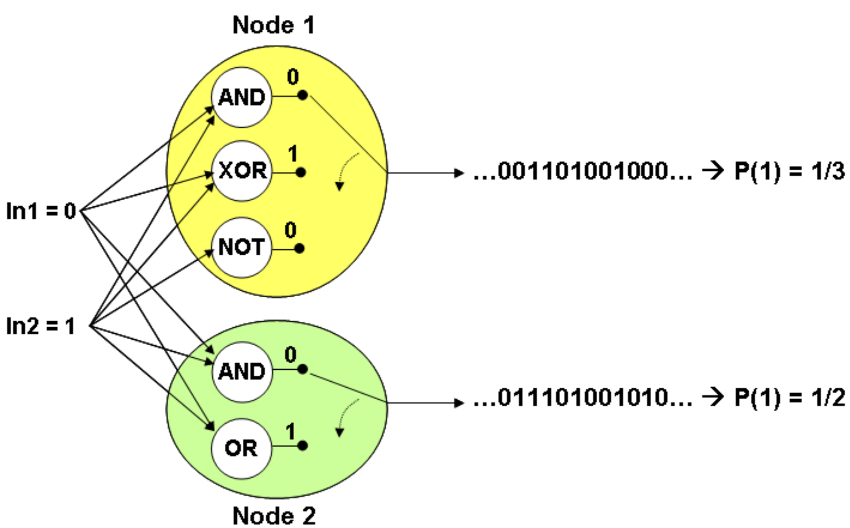
Connection between conditional probabilities in Bayesian networks and Boolean truth tables

Noisy Boolean transfer functions translate into conditional probabilities. The reverse is not true because Bayesian networks have more parameters.

in2	in1	P(out=0)	P(out=1)
0	0	1	0
0	1	p_1	$1 - p_1$
1	0	p_2	$1 - p_2$
1	1	$p_1 \times p_2$	$1 - p_1 \times p_2$

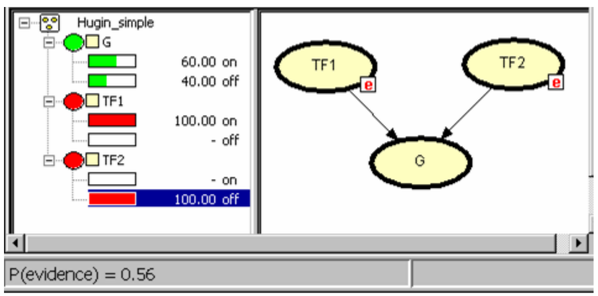
Noisy OR

A parallel of Bayesian conditional probabilities is a probabilistic Boolean network in which each node makes a probabilistic choice among several transfer functions.

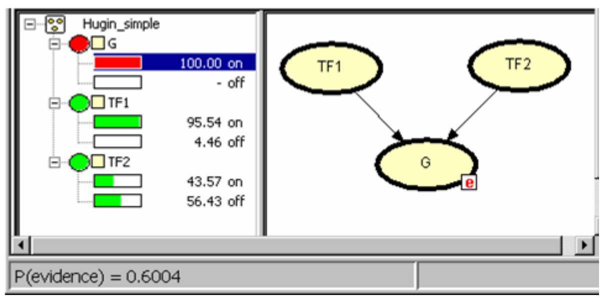


Bayesian model building and analysis is mainly computational

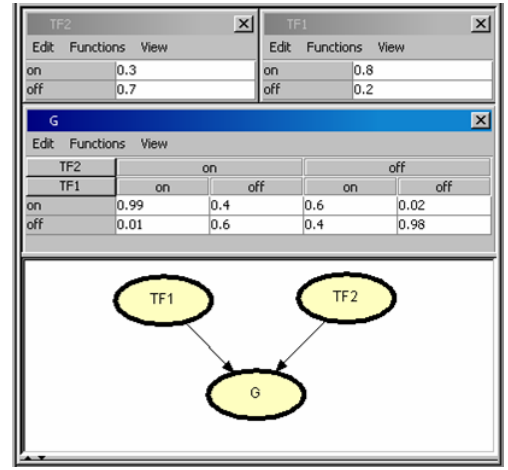
Implementations: Hugin (used in our textbook), Bayesian packages in R, Bayesian Network Toolbox for Matlab



Forward calculation:
TF1=on, TF2=off



Reverse calculation
(diagnostic): G=on



Hugin example input

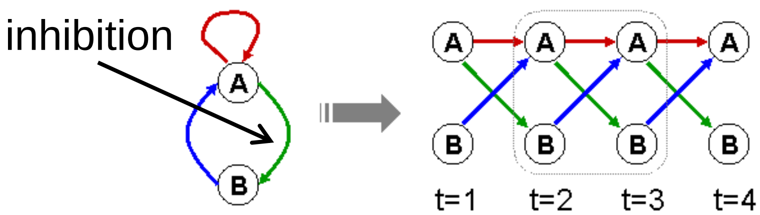
The most likely scenario in which G=on is not the same as the TF1/TF2 combination that gives the highest probability for G=on.

Q: why do you think this is so?

Modeling networks with feedback with Dynamic Bayesian Networks

Basic Bayesian networks are time-less, but feedback requires a time delay.

Idea: incorporate feedback by creating a copy of the network



A(t)	B(t)	P(A(t+1)=1
0	0	0.07
0	1	0.86
1	0	0.84
1	1	0.95

We need fine-resolution time-course data to learn the conditional probabilities.

Focused experiments, i.e. interventions, control are necessary to detect auto-regulatory feedback.

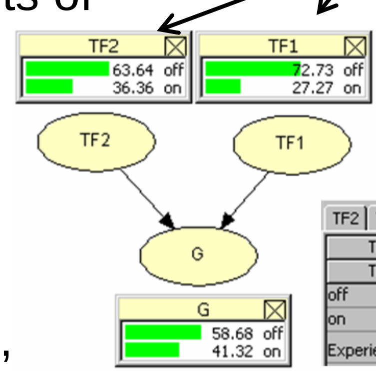
Constructing Bayesian models directly from data

- Each Bayesian network model needs two sets of parameters: the dependencies (network) and the values of the conditional dependency tables.

- If the network is known, the dependency tables can be inferred. If the network is not known, there is a search in the space of networks and each candidate network is evaluated after estimating its optimum conditional probability values.

G	TF1	TF2
off	off	off
off	off	off
off	off	off
on	off	off
on	off	on
on	off	on
off	off	on
off	off	on
on	on	off
off	on	off
on	on	off

- If for each node we have several measurements of node output value for each combination of input values, we can construct the conditional probability tables based on these measurements.



TF2 | TF1 | G

TF1	off		on	
TF2	off	on	off	on
off	0.75	0.5	0.333333	0.5
on	0.25	0.5	0.666667	0.5
Experience	4	4	3	0

- If there is no data for some input combinations, we can assume equal probability of each output for those inputs and then iteratively change to better fit the data.

Learning network structure from data

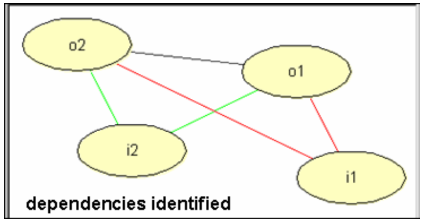
Most approaches generate an ensemble of plausible networks
The goodness score is defined based on Bayes’ theorem
$P(S|D) = P(D|S)P(S)/P(D)$

Estimate $P(D|S)$:
$\log P(D|S) = \log P(D|S, CPT) - (K/2) \log N$
CPT: estimated optimal probability tables,
K: # free parameters, N: # data samples

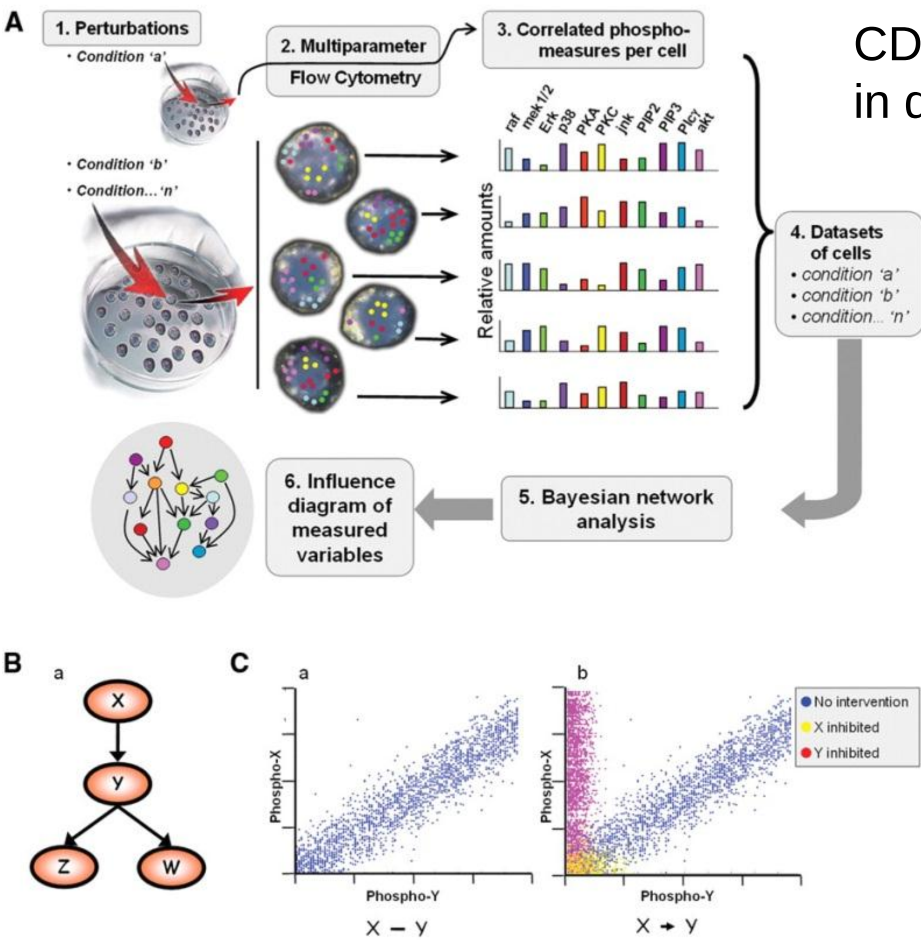
Maximize $\log P(D|S)$ over different networks.

One link of each color is sufficient to explain the data on top.
Edges that recur in many high-scoring networks are particularly plausible.

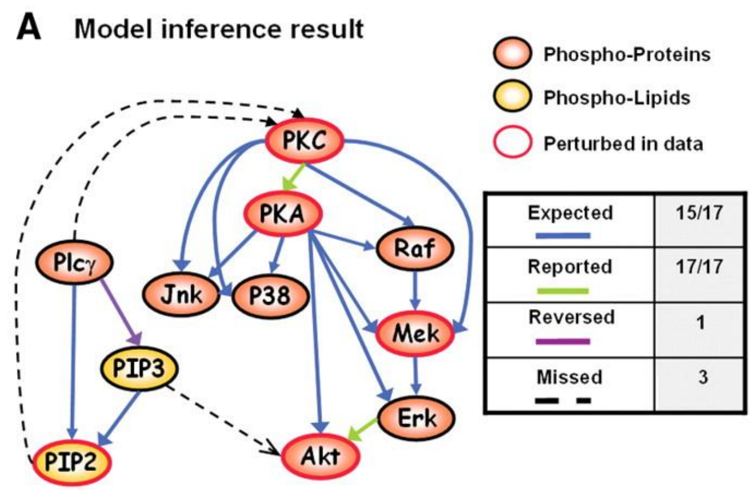
i1	i2	o1	o2
0	0	0	1
0	1	1	0
1	0	1	0
1	1	1	0
0	0	0	1
0	1	1	0
1	0	1	0
1	1	1	0
0	0	0	1
0	1	1	0
1	0	1	0
1	1	1	0
0	0	0	1
0	1	1	0
1	0	1	0
1	1	1	0
0	0	0	1
0	1	1	0
1	0	1	0
1	1	1	0



Example: inferring signaling networks from single cell data



CD4+ T cells, 11 proteins/lipids measured in different signaling and intervention cases.



correlation

causation can be inferred from interventions

Sachs et al Science 308, 523 (2005)

Current limitations of Bayesian networks

- Efficient network evaluation is absolutely essential in Bayesian network modeling.
- This is greatly facilitated by the two key assumptions of no feedback and no memory.
- But these assumptions also limit the utility of Bayesian networks as models of regulatory networks.
- A third limitation is the need of enough replicate data such that observation frequencies can be interpreted as conditional probabilities.